Exciton condensation and charge fractionalization in a topological insulator film

B. Seradjeh,[1] J.E. Moore,[2,3] and M. Franz[4]

[1]Department of Physics, University of Illinois, 1110 West Green St, Urbana, IL 61801-3080
[2]Department of Physics, University of California, Berkeley, CA 94720
[3]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[4]Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, Canada V6T 1Z1

(Dated: July 16, 2009)

An odd number of gapless Dirac fermions is guaranteed to exist at a surface of a strong topological insulator. We show that in a thin-film geometry and under external bias, electron-hole pairs that reside in these surface states can condense to form a novel exotic quantum state which we propose to call 'topological exciton condensate' (TEC). This TEC is similar in general terms to the exciton condensate recently argued to exist in a biased graphene bilayer, but with different topological properties. It exhibits a host of unusual properties including a stable zero mode and a fractional charge $\pm e/2$ carried by a singly quantized vortex in the TEC order parameter.

Introduction.—Recent advances in studies of band insulators with strong spin-orbit coupling revealed the existence of new topological invariants that characterize these materials [1]. Among the three-dimensional time-reversal ($\mathcal{T}$) invariant insulators, the most interesting phase implied by this classification is the "strong" topological insulator (STI), characterized by gapless fermionic states residing at its surface with an *odd* number of topologically protected nodes. These gapless states exhibit linear dispersion and behave as massless Dirac fermions familiar from the physics of graphene. Several real materials have been identified as STIs in pioneering experiments [2, 3] completed shortly after the theoretical predictions [4]. These rapid developments give hope that the new state of quantum matter realized in STIs might be relatively common in nature and raise the prospects of future practical applications.

The existence of an odd number of Dirac fermions leads to a number of exotic properties associated with surfaces of a STI. These include an exotic superconducting state induced by a proximity effect that supports Majorana fermions [5], a $\mathcal{T}$-breaking phase which exhibits a *fractional* quantum Hall effect [6], and an unusual 'axion' electromagnetic response [6, 7].

The wealth of exotic phenomena listed above stems from the possibility of inducing various types of mass terms in the otherwise massless Dirac fermion states at the surface of a STI. In this Letter we introduce and study a new type of mass gap that can be induced by a Coulomb interaction between the surface states of a thin STI film and can be characterized as a 'topological' exciton condensate (TEC). The idea is motivated by recent proposals to realize an exciton condensate in a symmetrically biased graphene bilayer [8, 9]. We argue below that TEC in an STI film might be more easily realized than in graphene and is a different, genuinely topological phase, distinguished by the presence of a zero-energy mode and fractional charge associated with its vortices.

Consider a film made of a STI placed inside a capacitor as in Fig. 1a. Imagine for simplicity that each surface harbors a single Dirac cone with the chemical potential μ initially tuned to the neutral point $\mu = 0$. When the capacitor is charged the Fermi levels in the two layers move in the opposite direction, creating a small electron Fermi surface in one layer and a small hole Fermi surface in the other. For arbitrarily weak repulsive interaction such a system will form an exciton condensate which may be pictured as a coherent liquid of electron-hole pairs residing in different layers.

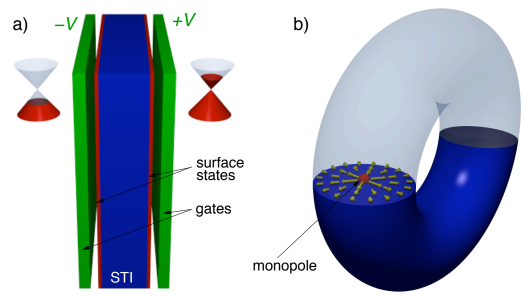

FIG. 1: (Color online) a) Schematic of the proposed device. b) The exciton condensate effectively joins the surfaces of the STI film resulting in toroidal topology. Arrows illustrate the magnetic field distribution of a planar monopole representing a vortex in the effective theory.

In what follows we use a simple model for the surface states to show how exciton condensation can be induced by the interlayer Coulomb interaction. By examining this model we then deduce some interesting properties of the underlying TEC. Specifically, we demonstrate that an isolated singly quantized vortex in the complex scalar order parameter characterizing TEC contains a zero mode and carries topologically protected exact fractional charge $\pm e/2$. We put our findings in the context of axion electrodynamics, the low-energy effective theory of STIs, and discuss prospects for experimental realization and detection of the predicted phenomena.

Model.—The gapless states associated with the two surfaces of the biased STI film can be described at low energies by a Dirac Hamiltonian [5]

$$H = \sum_{l=1,2} \psi_l^\dagger (v_l \boldsymbol{\sigma} \cdot \hat{\mathbf{p}} - \mu_l)\psi_l + U n_1 n_2, \qquad (1)$$

where $\psi_l = (c_{l\uparrow}, c_{l\downarrow})^T$ denotes the fermion spinor in surface layer l, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$ is the vector of Pauli matrices in the spin space, $\hat{\mathbf{p}} = -i\nabla$, and $v_l = (-1)^{l+1}v$ represents the Fermi velocity, assumed to be opposite at the two surfaces and $n_l = \psi_l^\dagger \psi_l$. We take $\mu_l = \mu + (-1)^l V$, with μ the intrinsic chemical potential and V the external bias. The last term in (1) describes the short-range part of the interlayer Coulomb potential, a sufficient minimal interaction for the formation of EC [10]. Also, we assume the film to be sufficiently thick so that any direct hopping of low-energy electrons between the surfaces can be neglected.

To describe the exciton condensation we decouple the interaction term in H using a matrix-valued order parameter $M = U\langle \psi_1 \psi_2^\dagger \rangle$. The expectation value is taken with respect to the mean-field Hamiltonian

$$H_{\rm MF} = H_0 + (\psi_1^\dagger M \psi_2 + {\rm h.c.}) + \frac{1}{U}{\rm Tr}(M^\dagger M), \qquad (2)$$

where H_0 denotes the kinetic term in (1). At this point it is useful to organize the Fermi fields into a single 4-component spinor $\Psi = (\psi_1, \psi_2)^T$. We can write $H_{\rm MF} = \Psi^\dagger \mathcal{H} \Psi + \frac{1}{U}{\rm Tr}(M^\dagger M)$ with a 4×4 matrix Hamiltonian

$$\mathcal{H} + \mu = \begin{pmatrix} v\boldsymbol{\sigma}\cdot\hat{\mathbf{p}} - V & M \\ M^\dagger & -v\boldsymbol{\sigma}\cdot\hat{\mathbf{p}} + V \end{pmatrix}. \qquad (3)$$

Various forms of matrix M describe different possibilities for the TEC order parameter. When μ is close to zero an order parameter that opens up a gap in the excitation spectrum will be favored because it leads to an overall reduction in kinetic energy. In the uniform system this will occur only for M diagonal in the spin space, i.e. $M = m\mathbb{1}$ with m a complex constant, since the part of $\mathcal{H}$ proportional to m then anticommutes with the kinetic term. The spectrum of $\mathcal{H}$ then contains four branches and reads

$$E_{\mathbf{k}\alpha s} = -\mu + \alpha\sqrt{(v|\mathbf{k}| + sV)^2 + |m|^2}, \quad \alpha, s = \pm 1. \quad (4)$$

Physically, this form of the matrix M implies non-zero expectation values $\langle c_{1\uparrow}c_{2\uparrow}^\dagger\rangle = \langle c_{1\downarrow}c_{2\downarrow}^\dagger\rangle = m/U$. We also note that this is the only choice of the order parameter that leaves the Hamiltonian $\mathcal{T}$-invariant.

Exciton condensate.—When $\mu = 0$ and $V \neq 0$ the exciton instability of our Hamiltonian (1) is formally equivalent to the Cooper instability in a metal and occurs at infinitesimal coupling U. To see what happens when μ is slightly detuned from zero, we consider the gap equation for m (now assumed to be real) obtained by minimizing the ground state energy $E_g = \sum'_{\mathbf{k}\alpha s} E_{\mathbf{k}\alpha s} + Nm^2/U$ with respect to m. This reads

$$\frac{m}{U} = -\frac{1}{2N}\sum_{\mathbf{k}\alpha s}{}' \frac{m}{E_{\mathbf{k}\alpha s} + \mu}, \qquad (5)$$

where the prime denotes a sum over the occupied states ($E_{\mathbf{k}\alpha s} < 0$) only. For general values of μ, V and U the gap equation must be solved numerically. One can, however, extract the value U_c of the critical coupling beyond which TEC is formed. In the experimentally relevant regime $\mu \ll V \ll \Lambda$, where Λ denotes the high-energy cutoff of the order of bandwidth, we find

$$U_c \simeq 4\Lambda\left(1 + \frac{V}{\Lambda}\ln\frac{V}{\mu}\right)^{-1}. \qquad (6)$$

The critical coupling is large, of the order of bandwidth, unless $\mu \ll Ve^{-\Lambda/V}$, in which case U_c becomes small and eventually reaches zero when $\mu \to 0$. In this limit the gap equation can be solved explicitly to obtain $m \approx 2\sqrt{V\Lambda}e^{-\Lambda^2/UV}$. In order to achieve exciton condensation for a given coupling strength U it is essential to tune μ as close to zero as possible and apply high bias voltage V. Henceforth we consider only the $\mu = 0$ situation.

Because of its exponential dependence on the coupling strength it is difficult to give a truly quantitative estimate of m and the relevant TEC transition temperature T_c for a realistic STI film. In the context of the graphene bilayer the estimates of T_c range from sub-Kelvin up to the room temperature, depending on the approximation employed [8, 9, 12]. Although we do not attempt such a quantitative analysis here we note that the situation in STI might be quite similar. On the one hand the intrinsic energy scales in known STIs are somewhat smaller than in graphene. On the other hand, screening is known to reduce the mean-field T_c by a factor $\sim e^{\mathcal{N}}$, where $\mathcal{N}$ denotes the number of surface Dirac modes. Formation of TEC in an STI film made from Bi_2Se_3 will be therefore aided by the fact that this material exhibits $\mathcal{N} = 1$ [2] compared to $\mathcal{N} = 4$ in graphene (due to valley and spin degeneracies).

Vortex zero modes.—In the following we adopt the point of view that, based on the above analysis, formation of TEC is likely to occur under experimentally achievable conditions and focus on its unique properties. To this end it is convenient to write the Hamiltonian (3) in a more customary form using the Dirac matrices in the Weyl representation, $\gamma_j = i\tau_2 \otimes \sigma_j$, $j = 1, 2, 3$, $\gamma_0 = \tau_1 \otimes \mathbb{1}$ and $\gamma_5 = -i\gamma_0\gamma_1\gamma_2\gamma_3 = \tau_3 \otimes \mathbb{1}$, where τ_j are Pauli matrices in the layer space. We obtain

$$\mathcal{H} = \gamma_0\left(\gamma_1\hat{p}_x + \gamma_2\hat{p}_y + V\gamma_0\gamma_5 + |m|e^{-{\rm i}\gamma_5\chi}\right), \qquad (7)$$

where we have set $v = 1$ and used a polar representation $m = |m|e^{i\chi}$ of the complex TEC order parameter.

The Hamiltonian (7) formally coincides with the one used to describe the effect of EC near one of the valleys of the biased graphene bilayer system [10]. That work established existence of an *exact zero mode* of the Hamiltonian (7) in the presence of a singly quantized vortex in the EC order parameter $m = m_0 e^{i\varphi}$, with φ the polar angle. The zero-energy eigenstate has the form $\Psi_0 = (f, g, ig^*, -if^*)^T$ with $f = Ae^{-m_0 r}J_0(Vr)$, $g = -iAe^{i\theta}e^{-m_0 r}J_1(Vr)$ and A the normalization constant. In graphene, valleys always come in pairs. The

fermionic zero modes are thus doubled and split due to intervalley scattering. Because of this, no exact zero modes survive in the graphene bilayer.

In a STI, by contrast, there is always an *odd number* of valleys associated with the surface [1]. When $\mathcal{N}=1$, as in Bi_2Se_3, the zero mode attached to a singly quantized vortex will remain exact, as long as the vortex stays well separated from other vortices or system edges. This finding, along with the fractional charge discussed below, constitutes the key universal difference between the STI exciton condensate and other proposed exciton condensates and is the main result of this work.

For general odd $\mathcal{N}>1$ we expect $\mathcal{N}-1$ zero modes to split symmetrically around zero energy while the remaining zero mode will persist. This conclusion follows from the property $\gamma_2\mathcal{H}^*\gamma_2=\mathcal{H}$ which together with $\gamma_2^2=-1$ implies the spectral symmetry around the zero energy of eigenstates of $\mathcal{H}$. Specifically, for each eigenstate Ψ_E of energy E there exists an eigenstate $\Omega\Psi_E$ with energy $-E$. Here $\Omega=\gamma_2 K$ is an antiunitary operator and K denotes complex conjugation. It also holds that $\{\Omega,\mathcal{H}\}=0$.

Fractional charge.—A localized zero mode in a particle-hole symmetric system is known to carry a fractional charge $\pm e/2$ [13, 14, 15, 16]. Thus, we expect our vortices to be fractionally charged. To see how this occurs in the present system and also to deduce some of its other interesting properties, let us consider an operator O, represented by a constant 4×4 Hermitian matrix, acting in the space of wavefunctions $\Psi_E(\mathbf{r})$. Following [11] consider now the quantity

$$\mathcal{O}\equiv\sum_E\langle\Psi_E|O|\Psi_E\rangle=\frac{1}{4}N\mathrm{tr}(O), \tag{8}$$

where N is the total number of quantum states in a suitably regularized system (e.g. on the lattice and in finite volume). The last equality in (8) follows from the completeness of states. We may write

$$\mathcal{O}=\left(\sum_{E<0}+\sum_{E>0}\right)\langle\Psi_E|O|\Psi_E\rangle+\langle\Psi_0|O|\Psi_0\rangle \tag{9}$$

The spectral symmetry generated by Ω and its antiunitarity, expressed as $\langle\Omega\Psi_1|\Omega\Psi_2\rangle=\langle\Psi_1|\Psi_2\rangle^*$, imply $\sum_{E>0}\langle\Psi_E|O|\Psi_E\rangle=\sum_{E<0}\langle\Psi_E|(\Omega^{-1}O\Omega)^\dagger|\Psi_E\rangle$. If O furthermore *commutes* with Ω then the last term becomes simply $\sum_{E<0}\langle\Psi_E|O|\Psi_E\rangle$ and we can combine Eqs. (8-9) to obtain

$$\sum_{E<0}\langle\Psi_E|O|\Psi_E\rangle=\frac{1}{2}\left[\frac{N}{4}\mathrm{tr}(O)-\langle\Psi_0|O|\Psi_0\rangle\right]. \tag{10}$$

The expectation value of an observable represented by a constant 4×4 matrix that commutes with Ω, taken over all occupied negative-energy eigenstates of $\mathcal{H}$, is determined solely by the value of $\mathrm{tr}(O)$ and the zero-mode eigenstate of $\mathcal{H}$. For an infinite system in continuum Eq. (10) will be useful for quantities independent of N; this occurs when O is traceless or else for quantities represented as differences so that $N\mathrm{tr}(O)$ drops out. In such cases the expectation value only depends on the zero mode and its value is expected to be robust. Specific examples follow below.

The charge operator is represented by a 4×4 unit matrix $O_Q=e\mathbb{1}$. The charge bound to a vortex can be expressed as

$$Q_V=e\sum_{E<0}\big(\langle\Psi_E|\mathbb{1}|\Psi_E\rangle_1-\langle\Psi_E|\mathbb{1}|\Psi_E\rangle_0\big), \tag{11}$$

where subscripts 1 and 0 refer to the state with one and zero vortices respectively. Using Eq. (10) we find $Q_V=-\frac{e}{2}\langle\Psi_0|\Psi_0\rangle_1=-\frac{e}{2}$, as expected. We note that Eq. (11) assumed the zero mode to be unoccupied; if we occupy it by an electron then the vortex charge becomes $+e/2$.

Other quantities of interest include the spin $\mathbf{S}$ and the axial charge Q^5 carried by the vortex, defined as the charge difference between the layers. These are represented by matrices $O_{\mathbf{S}}=\frac{1}{2}\gamma_0\vec{\gamma}\gamma_5$ and $O_{Q^5}=e\gamma_5$. Unfortunately these anticommute with Ω making Eq. (10) inapplicable. A quantity that can be calculated is the interlayer spin polarization $\Delta\mathbf{S}$, represented by $O_{\Delta\mathbf{S}}=\frac{1}{2}\gamma_0\vec{\gamma}$. A straightforward calculation shows that, in the presence of a vortex, $\langle\Delta S_x\rangle=\langle\Delta S_y\rangle=0$ while $\langle\Delta S_z\rangle$ varies smoothly from $\frac{1}{2}$ to 0 as we tune V/m from 0 to infinity. This implies that the vortex carries a fractional value of spin polarization between the layers. We note, however, that due to the entanglement of the spin and momenta in the STI, the total spin operators in the TEC or the STI surfaces are not sharp.

The symmetry generated by Ω is a combination of $\mathcal{T}$ and spatial parity $\mathcal{P}$. The latter will be broken in the presence of non-magnetic impurities and the zero mode will no longer be exact. However, the following general argument shows that the fractional charge remains precisely quantized as long as the bulk is gapped and $\mathcal{T}$ is preserved.

Axion electrodynamics.— As demonstrated in Refs. [6, 7] the response of STI to external electromagnetic field is that of an 'axion' medium [17] and can be mathematically implemented by adding a term

$$\Delta\mathcal{L}_{\mathrm{axion}}=\frac{\theta}{2\pi}\frac{e^2}{hc}\mathbf{B}\cdot\mathbf{E} \tag{12}$$

to the usual Maxwell Lagrangian. An ordinary insulator has $\theta=0$ while the STI exhibits $\theta=\pi$, the two values permitted by the time-reversal symmetry. When a surface of a STI is gapped by a $\mathcal{T}$-breaking perturbation, such as an applied magnetic field, θ varies smoothly between π and 0. As a result the surface behaves as a quantum Hall fluid [6, 17].

As noted above, TEC does not break $\mathcal{T}$. Instead, the TEC order parameter effectively identifies the opposite surfaces of the film by allowing, at the mean-field level of the Hamiltonian (2), the electrons to hop between them. Dirac fermions at the surfaces are gapped without violating $\mathcal{T}$ by a perturbation that effectively removes the

surfaces as illustrated in Fig. 1b. Because the TEC order parameter is in general complex the electrons hopping between the surfaces may acquire a nontrivial phase. This is easily included by postulating twisted periodic boundary conditions with a twist equal to the phase of the TEC order parameter χ. The situation is easiest to visualize in the context of a lattice model of STI where electrons traversing the bonds that connect the two surfaces acquire a phase χ. This extra phase can be also viewed as resulting from an electromagnetic vector potential $\delta\mathbf{A} = \hat{z}\chi(\Phi_0/a)$ localized in the layer between the surfaces, with $\Phi_0 = hc/e$ the flux quantum and a the lattice spacing.

In a vortex configuration with $\chi(\mathbf{r}) = \varphi$ it is easy to see that $\delta\mathbf{A}(\mathbf{r})$ has the form of a planar *monopole*: the magnetic field $\delta\mathbf{B}$ radiates outward from the vortex center in the plane of the surface as illustrated in Fig. 1b. The total flux is Φ_0. This vector potential is fictitious in the sense that $\delta\mathbf{B}$ cannot be detected by an outside probe. To electrons, however, $\delta\mathbf{A}$ is indistinguishable from the real vector potential and must be included in Eq. (12) when evaluating the response of the STI. A unit magnetic monopole in a 'θ-vacuum' described by Eq. (12) is known to carry electric charge $-e(\theta/2\pi + n)$ with n integer [18]. Applied to TEC this gives vortex charge $-e(1/2+n)$, consistent with our finding of the fractional charge $\pm e/2$ bound to the vortex.

Outlook and open questions.— Recent advances in materials engineering and fabrication give hope that the exotic state of matter identified in this work can be achieved and probed in the near future. From an experimental point of view, there are several significant advantages of the proposed phase. First, we believe that making an exciton condensate between the surfaces of a single film, rather than from two 2D materials (e.g., graphene) with an insulator between them, is easier because it does not require creating a pinhole-free insulating layer with defect-free junctions to the two 2D materials. Second, attaching leads to a surface of a film should be considerably easier than attaching leads to graphene. Once the leads are in place it should be straightforward to identify the onset of the exciton condensation in a transport measurement as demonstrated in recent literature [19]. Third, we note that the existence of the zero-energy Dirac fermion can be detected in the same way as a standard midgap impurity state by optical methods or in a careful transport measurement. In the TEC phase a transport measurement will reveal a gap at low temperature. As the temperature nears the transition temperature, conduction becomes dominated by the charges bound to vortices and is proportional to the number of thermally excited vortices. Fractional charge can be probed, at least in principle, by the shot noise analysis of resistivity [20].

From a theoretical point of view TEC in STI film is fundamentally interesting for several reasons. To our knowledge, TEC is the first example of a new symmetry-breaking phase enabled by the special properties of topological insulators. It differs from the superconducting state generated at the surface by proximity effect [5] because in that case there is no new symmetry breaking. It differs from the ordinary exciton condensate, which is in the same universality class as a ^{4}He film, because the zero mode attached to a vortex is stable. In the case of graphene bilayer the zero modes are not protected due to inter-valley scattering. In this respect as in several others, topological insulators allow the realization of physics that is spoiled in graphene by inter-valley scattering. The consequence is a distinct low-temperature phase of matter with fractionally charged topological excitations whose exchange statistics presents an interesting open question.

Acknowledgment.—Support for this work came from NSERC, CIfAR, NSF, ICMT at UIUC and the Killam Foundation. The authors also acknowledge The Banff International Research Station where this collaboration was initiated.

[1] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98** 106803 (2007); J. E. Moore and L. Balents, Phys. Rev. B **75** 121306(R) (2007); R. Roy, arXiv:cond-mat/0607531.

[2] D. Hsieh *et al.*, Nature **452**, 970 (2008).

[3] Y. Xia *et al.*, arXiv:0812.2078.

[4] J. C. Y. Teo, L. Fu, and C. L. Kane, Phys. Rev. B **78** 045426 (2008); H.-J. Zhang *et al.*, arXiv:0901.2762.

[5] L. Fu and C. L. Kane, Phys. Rev. Lett. **100**, 096407 (2008).

[6] X.-L. Qi, T. Hughes, and S.-C. Zhang, Phys. Rev. B **78**, 195424 (2008).

[7] A.M. Essin, J.E. Moore, and D. Vanderbilt, Phys. Rev. Lett. **102**, 146805 (2009).

[8] Y. E. Lozovik and A. A. Sokolik, JETP Lett. **87**, 55 (2008).

[9] H. Min, R. Bistritzer, J.-J. Su, and A. H. MacDonald, Phys. Rev. B **78**, 121401(R) (2008); see also arXiv:0810.0331 for comment on Ref. [12].

[10] B. Seradjeh, H. Weber, and M. Franz, Phys. Rev. Lett. **101**, 246404 (2008).

[11] I. F. Herbut, Phys. Rev. Lett. **99**, 206404 (2007).

[12] M. Y. Kharitonov and K. B. Efetov, Phys. Rev. B **78**, 241401(R) (2008).

[13] R. Jackiw and C. Rebbi, Phys. Rev. D **13**, 3398 (1976).

[14] J. Goldstone and F. Wilczek, Phys. Rev. Lett. **47**, 986 (1981).

[15] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).

[16] C.-Y. Hou, C. Chamon, and C. Mudry, Phys. Rev. Lett. **98**, 186809 (2007).

[17] F. Wilczek, Phys. Rev. Lett. **58**, 1799 (1987).

[18] E. Witten, Phys. Lett. B**86**, 283 (1979).

[19] J. P. Eisenstein and A. H. MacDonald, Nature **432**, 691 (2004).

[20] L. Saminadayar, D.C. Glattli, Y. Jin, and B. Etienne, Phys. Rev. Lett. **79**, 2526 (1997).